

# APPLICATION UNDER UNITED STATES PATENT LAWS

Atty. Dkt. No. PW 280336  
(M#)

Invention: COALESCING INFORMATION FROM MULTIPLE SOURCES BASED ON PRIORITY RULES

Inventor (s): Dale T. TAYLOR  
Dave L. METHENY

Pillsbury Winthrop LLP  
Intellectual Property Group  
1600 Tysons Boulevard

McLean, VA 22102  
Attorneys  
Telephone: (703) 905-2000

## This is a:

- ☐ Provisional Application
- ☒ Regular Utility Application
- ☐ Continuing Application
  - ☐ The contents of the parent are incorporated by reference
- ☐ PCT National Phase Application
- ☐ Design Application
- ☐ Reissue Application
- ☐ Plant Application
- ☐ Substitute Specification
  - Sub. Spec Filed \_\_\_\_\_
  - in App. No. \_\_\_\_\_ / \_\_\_\_\_
- ☐ Marked up Specification re
  - Sub. Spec. filed \_\_\_\_\_
  - In App. No. \_\_\_\_\_ / \_\_\_\_\_

## SPECIFICATION

09290" 52216850

## **COALESCING INFORMATION FROM MULTIPLE SOURCES BASED ON PRIORITY RULES**

### **Reservation of Copyright**

**[0001]** This patent document contains information subject to copyright protection.

The copyright owner has no objection to the facsimile reproduction by anyone of the patent document or the patent, as it appears in the U.S. Patent and Trademark Office files or records but otherwise reserves all copyright rights whatsoever.

### **BACKGROUND**

**[0002]** Aspects of the present invention relate to gathering and use of information.

Other aspects of the present invention relate to information synchronization.

In a society that is becoming increasingly networked, information is often gathered from different sources across a network and used to make various decisions. For example, managers of computer networks may use discovery agents to gather information from devices in the network. Examples of the information collected by discovery agents include operational status (e.g., the availability), transmission facilities (e.g., the bandwidth that a device can handle) of the underlying devices, or the capabilities of the installed devices which may include software and available hardware. Such collected information may be used to evaluate network performance such as determining the bottleneck of the network traffic, to determine the structure and the operation of the network, and to improve the network performance by reconfiguring the network based on the gathered information.

[0003] In a different application, database administrators may deploy similar agents to monitor data manipulation acts performed on devices that host different copies of the same data or to identify different versions of the same file stored on different devices. The information gathered from these monitoring agents may be used to determine appropriate consolidation operations to enforce data integrity. For example, a user may have a file stored on different devices with the same name (e.g., on a personal computer at home and on a Palm Pilot). The files with the same name may correspond to different versions of the same file that is edited under different circumstances (e.g., the user may have edited the file on the personal computer on Monday and edited it again while the user is on the road on Tuesday). An automated data administrator may automatically consolidate these different versions to maintain the data integrity of the file.

[0004] The aggregated information gathered from multiple sources (via or through agents) may contain duplicate data. For example, when a network manager sends inquiries to multiple agents to search for an available router device, more than one agent may respond (e.g., multiple routers are available). The inquiry from a database administrator to identify a file on different devices may yield multiple responses.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

[0005] The present invention is further described in terms of exemplary embodiments, which will be described in detail with reference to the drawings. These embodiments are non-limiting exemplary embodiments, in which like reference numerals represent similar parts throughout the several views of the drawings, and wherein:

[0006] Fig. 1 depicts a high level architecture of embodiments of the present invention;

[0007] Fig. 2 depicts the internal structure of an agent interacting with a device group and a priority rule-based coalescing mechanism according to an embodiment of the present invention;

[0008] Fig. 3 depicts the internal structure of a priority rule-based coalescing mechanism according to embodiments of the present invention;

[0009] Fig. 4 shows an exemplary construct of priority rules according to an embodiment of the present invention;

[0010] Fig. 5 is an exemplary flowchart of a process, in which information from different sources is coalesced, in a synchronous mode, based on priority rules according to an embodiment of the present invention;

[0011] Fig. 6 is an exemplary flowchart of a process, in which an agent supplies information to a priority rule-based coalescing mechanism in a synchronous mode, according to an embodiment of the present invention;

[0012] Fig. 7 is an exemplary flowchart of a process, in which information from different sources is coalesced, in an asynchronous mode, based on priority rules according to a different embodiment of the present invention; and

[0013] Fig. 8 is an exemplary flowchart of a process, in which an agent supplies information to a priority rule-based coalescing mechanism in an asynchronous mode, according to a different embodiment of the present invention.

## DETAILED DESCRIPTION

[0014] The inventions are described below, with reference to detailed illustrative embodiments. The inventions can be embodied in a wide variety of forms, some of which may be quite different from those of the disclosed exemplary embodiments. Consequently, the specific structural and functional detail is disclosed herein are merely representative and do not limit the scope of the invention.

[0015] The processing described below may be performed by a general-purpose computer alone or in connection with a special purpose computer. Such processing may be performed by a single platform or by a distributed processing platform. In addition, such processing and functionality can be implemented in the form of special purpose hardware or in the form of software being run by a general-purpose computer. Any data handled in such processing or created as a result of such processing can be stored in any memory as is conventional in the art. By way of example, such data may be stored in a temporary memory, such as in the RAM of a given computer system or subsystem. In addition, or in the alternative, such data may be stored in longer-term storage devices, for example, magnetic disks, rewritable optical disks, and so on. For purposes of the disclosure herein, a computer-readable media may comprise any form of data storage mechanism, including such existing memory technologies as well as hardware or circuit representations of such structures and of such data.

[0016] Fig. 1 depicts a high level architecture of embodiments of the present invention. In Fig. 1, a configuration 100, consistent with the present invention, comprises a plurality of n agents (agent 1, 135, ..., agent i, 145, agent j, 155, ..., agent n, 165) associated with a plurality of n corresponding device groups (device group 1, 130, ..., device group i,

140, device group j, 150,..., device group n, 160) and a priority rule-based coalescing mechanism 110 connecting with the plurality of n agents via a network 120. The priority rule-based coalescing mechanism 110 in Fig. 1 receives information, related to the devices in the device groups, from the corresponding agents and coalesces the received information according to pre-defined priority rules.

[0017] Each of the device groups may include one or more devices (not shown in Fig. 1). Such a device represents a generic device that may be a computer, a disk drive, a router, a file system, or a home appliance. The devices in a single device group may all reside in a same Local Area Network (LAN) or they may be distributed in a Wide Area Network (WAN). The agent associated with a device group may communicate with the associated devices within a LAN, a WAN, or the Internet (not shown in Fig. 1). The network 120 in Fig. 1 connecting the agents (135,...,145,155,...,165) and the priority rule-based coalescing mechanism 110 represents a generic network such as the Internet.

[0018] An agent connects to and communicates with the devices in its corresponding device group to, for example, acquire status information of the devices and to send such information to the priority rule-based coalescing mechanism 110 via the network 120. These devices may operate on different platforms (e.g., Unix platform and PC platform) or use different protocols (e.g., a computer and a home appliance such as a Video Cassette Recorder (VCR) may use different protocols) to communicate with their agent.

[0019] In configuration 100, an agent is capable of communicating with the devices in its device group that are running on different platforms and using different protocols. An agent serves as an interface between the devices in its device group and the priority rule-based coalescing mechanism 110. While interacting with the priority rule-based coalescing

mechanism 110, a uniform schema or convention may be adopted. For example, a schema defined using the eXtensible Markup Language (XML) may be pre-specified to deliver the information from agents to the priority rule-based coalescing mechanism 110.

[0020] A device may send information as a response to its corresponding agent based on an inquiry. For example, an agent may query the status of a disk drive (a device) in its associated device group. An agent may query the timestamp of a particular file in a file system (a different device). A device may also send information to its associated agent based on its own initiative. For example, when a device finishes its current job and ready for other tasks, the device may actively send information to its associated agent to advertise its availability.

[0021] The priority rule-based coalescing mechanism 110 receives information from multiple sources (agents) and coalesces the received information based on certain priority rules. Such priority rules may specify the relative importance of different agents. The relative importance may be defined with respect to some criteria that may be specified according to application needs. A default priority may be assumed for an agent if no priority is defined explicitly. Priority rules provide a base for the priority rule-based coalescing mechanism 110 to prioritize the information from different sources.

[0022] In addition to the priority rules, the priority rule-based coalescing mechanism 110 may also integrate other types of information during coalescing. For example, a coalescing decision about the information relevant to a file may be made based on both the priority of the corresponding file manager (a device) as well as the timestamp of the file. For instance, if two copies (e.g., corresponding to different versions) of the same file (stored on different devices) have the same priority but one has a more recent timestamp, the one with

the more recent timestamp may be kept during the coalescing and the other may be removed during the coalescing.

**[0023]** Priorities may also be defined with respect to devices in a device group. For example, a Central Processing Unit (CPU) of a computer in a device group, represented by an agent, may have a higher priority than a CPU of another computer in a different device group, represented by a different agent. The priority of a device in a device group may be inherited from the priority assigned to its associated agent or may be defined explicitly. When an explicit priority is assigned to a device, it may override the inherited priority of the device.

**[0024]** Fig. 2 depicts the internal structure of an agent (e.g., agent i, 145) interacting with a device group and the priority rule-based coalescing mechanism 110 according to an embodiment of the present invention. In Fig. 2, the agent i, 145, is connected to the device group i, 140, which comprises a plurality of devices, device 1, 210a, ..., device i, 210b, ..., device m, 210c. The agent 145 comprises a device interface 220, an information formatting mechanism 230, and a network communication mechanism 240. The device interface 220 is responsible for interacting with the devices (210a, ..., 210b, ..., 210c) in the device group 140 to send inquiries to and to receive information from the devices.

**[0025]** The information formatting mechanism 230 takes the information received from the devices and organizes it according to a pre-defined convention prior to sending the information to the priority rule-based coalescing mechanism 110. For example, if a file system sends a timestamp of a particular file together with the file name to the device interface (which may be subsequently sent to the priority rule-based coalescing mechanism 110), the information formatting mechanism 230 may organize the information according to a schema pre-specified in XML. The schema may specify tags for different types of



information and how these tags are arranged. For example, according to a schema, a timestamp and the file name are placed into separate tag fields. Such a schema may be designed so that both the priority rule-based coalescing mechanism 110 and the agents 135, ..., 145, 155, ..., 165 can understand the information.

[0026] The formatted information generated by the information formatting mechanism 230, is sent to the priority rule-based coalescing mechanism 110 via the network communication mechanism 240. The network communication mechanism 240 is capable of interacting with the priority rule-based coalescing mechanism 110 using a protocol that is suitable for the network 120.

[0027] Fig. 3 depicts the internal structure of the priority rule-based coalescing mechanism 110 interacting with an agent (e.g., agent 145) according to embodiments of the present invention. The priority rule-based coalescing mechanism 110 includes a network communication mechanism 310, a transformation mechanism 320, a coalescing mechanism 330, a priority rule set-up mechanism 360, a coalesced file initialization mechanism 350, a priority rule access mechanism 370, a parsing mechanism 380, a priority rule database 365, and a coalesced file 335.

[0028] The responsibility of the priority rule-based coalescing mechanism 110 is to synchronize information from multiple sources. For example, different sources may send duplicate information or variants of the same information. Such information may have to be synchronized before it can be effectively used. The synchronization or coalescing may include removing duplicates or choosing one piece of specific information among a plurality of variants according to some criterion (which may be application dependent). According to the present invention, criteria used to coalesce the information from multiple agents include

priorities associated with different sources of information (agents). Such priorities are centrally set-up and controlled within the priority rule-based coalescing mechanism 110. That is, sources of the information (i.e., agents) send information without priority information or without concerning about the specifics of how the information will be processed or whether the information will be used (or kept).

[0029] In addition to the priorities, the priority rule-based coalescing mechanism 110 may also combine other types of information (e.g., parameters from devices) with the priorities during coalescing. For instance, the bandwidth of a router may be used, together with the priority of the router, to make a coalescing decision.

[0030] The network communication mechanism 310 is responsible for interacting with the agents (e.g., agent i, 145), including sending inquiries and receiving information from the agents. The interaction may be carried out through a pre-defined protocol or convention. For example, the information sent from the agents may be constructed according to a schema described using a standard language such as XML.

[0031] The priority rule set-up mechanism 360 initializes or updates the priority rules and stores the priority rules in the priority rule database 365. The priority rules may define the relative importance of both agents and devices and they may be specified in terms of certain structure expressed in an understood syntax.

[0032] Fig. 4 shows an exemplary construct of priority rules according to an embodiment of the present invention. Priority rules 410 stored in the priority rule database 365 comprise the priority rules for individual agents (agent 1 priority rules 420, agent 2 priority rules 430, ..., agent n priority rules 440). The priority rules for a particular agent (e.g., agent 2 priority rules 430) may define both a priority for the agent and individual priority

rules for different devices that are in the device group of the corresponding agent. For example, in Fig. 4, the agent 2 priority rules include device 1 priority rule 450,..., device m priority rule 460.

[0033] The structure of the priority rules may be defined as a schema specified in certain language. For example, XML may be used to specify the construct of priority rules. Below is an exemplary construct for priority rules specified in XML:

```
<priority>
    <agentname priority="priority-value">
        <device1-tag> "priority-value" </device1-tag>
        <device2-tag> "priority-value" </device2-tag>
        :
        <deviceM-tag> "priority-value" </deviceM-tag>
    </agentname>
    :
    :
</priority>
```

where <priority> ... </priority> encloses the priority definition block which may include priority rule definitions for a plurality of agents, "agentname" corresponds to the identification of an agent, "priority" is a keyword denoting where the priority should be defined, "priority-value" may be a numeral defining a priority level, and </agentname> ends the definition of the priority for the agent named "agentname". According to the exemplary schema, device priorities may be defined inside the agent priority definition block (enclosed in <agentname> </agentname>). In the above exemplary schema, "device-tag" identifies a

device and the “priority-value” is a numeral defining the priority of the device named in “device-tag”. Tag </device-tag> closes the priority definition block for the device. Based on the above exemplary schema, the following example shows priority rules defined according to the exemplary schema:

```
<priority>
  <agent1 priority="10">
    <host> 2 </host>
    <URL> 4 </URL>
    <computer> 0 </computer>
  </agent1>
  <agent2>
    <host> 5 </host>
    <IPAddress> 3 </IPAddress>
    <MACAddress> 12 </MACAddress>
    <route> 10 </route>
  </agent2>
</priority>
```

**[0034]** In the above example, the priority for agent 1 is defined as level 10 and the priority of agent 2 is assumed default priority level, which may be pre-defined as, for example, level 0. A larger number may be defined to represent a higher priority. The devices associated with an agent may default to inherit the priority level assigned to their associated agent except when it is explicitly defined otherwise to override the inherited priority.

**[0035]** Devices associated with agent1 in the above example have, by default, precedence over any other device associated with other agents (all other agents in the above example have default priority level 0) except when it is explicitly defined differently. For instance, device “y” of agent1 has a higher priority (level 10 – because it is not defined so it uses the default) than that of device “y” associated with agent2 (level 0). However, agent2’s “MACaddress” has a higher priority than any other agent’s “MACaddress” (because they are all default to level 0 – except for agent1’s which defaults to 10). Furthermore, in the above example, device “host” generated by agent2, if any, has a priority of 5 which overrides any “host” generated by any other agent including agent1 because it has been downgraded to level 2. The “route” generated by either agent1 or agent2 have the same level of priority (level 10). Depending on the specific implementation of the coalescing mechanism, information from agents with the same level of priority may be both kept in the final coalesced file. A coalescing mechanism may also be implemented to remove one using other criteria. Examples of other criteria may include timestamp or bandwidth. For instance, if two “routes” have the same priority but one allows higher bandwidth, the one with the higher bandwidth may be kept and the other may be removed from the coalesced file.

**[0036]** Referring back to Fig. 3, the role of the priority rule set-up mechanism 360 is to set up or update priority rules with respect to agents and the devices in their associated device groups. Priority rules may be updated when needed after the initial set-up. Such update may include adding new priority rules when new agents or new devices are added or deleting priority rules when the corresponding agents or devices are removed. Initial priority may also be changed which may either upgrades or downgrades the initial priority level.

[0037] Priority rules are stored in the priority rule database 365. The priority rule access mechanism 370 retrieves priority rules whenever it is needed. For example, it may be triggered to retrieve relevant priority rules when new information is received by the network communication mechanism 310 from an agent. Based on received information (e.g., from a particular agent), the network communication mechanism 310 may instruct the priority rule access mechanism 370 to retrieve the rules associated with the agent and its associated devices. Such retrieved priority rules may be fed to both the transformation mechanism 320 (explained later) and the parsing mechanism 380 to parse the rules before they can be used by the coalescing mechanism 330. Other embodiments (not shown in Fig. 3) may also be implemented in which the retrieved priority rules may be directly used by both the transformation mechanism 320 and the coalescing mechanism 330 without being parsed first.

[0038] According to the present invention, the agents do not concern (or have no knowledge) about the priority associated with them. The information sent from the agents does not include priority information. When information is received from an agent, associated priority information (retrieved by the priority rule access mechanism 370) may be incorporated into the received information before the information is coalesced. In Fig. 3, the transformation mechanism 320 takes the information received by the network communication mechanism 310 (from an agent) and the priority rules retrieved by the priority rule access mechanism 370 as input and generates modified information, in which each piece of information is coupled with appropriate priorities. For instance, if following information is received from an agent named "agent1":

<host> morgan </host>

<IPAddress> 128.34.56.78.9 </IPAddress>.

According to the priority rule for agent1 (discussed earlier), the agent1 has a priority level 10 and its associated "host" and "IPAddress" have priority levels 2 and 10, respectively. In this case, the transformation mechanism 320 may generate modified information as follows:

<host \_pri \_ = "2"> morgan </host>

<IPAddress \_pri \_ = "10"> 128.34.56.78.9 </IPAddress>.

[0039] Such modified information is fed to the coalescing mechanism 330. With appropriate priority information attached to the original information, the coalescing mechanism 330 coalesces the information accordingly and generates a coalesced file, which contains information that has been properly coalesced or synchronized. For example, duplicate information from multiple agents may be removed according to the priority rules.

[0040] As discussed earlier, the removal of a piece of information from a particular device may also be determined based on the operational parameters of the device. For example, if two devices have the same priority level but one has a higher computational power, the coalescing mechanism 330 may combine the priority with the computational power of the device and decides to retain the piece of information from the device with higher computational power. The coalesced file may be initially set up by the coalesced file initialization mechanism 350. The final coalesced file may be stored in the coalesced file database 335.

[0041] The coalescing mechanism 330 may perform coalescing in a different mode. One may be a synchronous mode and the other may be an asynchronous mode. In a synchronous mode, the coalescing mechanism 330 actively initiates each act in the coalescing process by sending an inquiry to an agent to collect information, coalescing the information, generating an updated coalesced file, and then moving on to the next agent. All is done in a

synchronous fashion. In an asynchronous mode, the coalescing process is activated whenever information is received from an agent. That is, the coalescing process is activated by an agent.

[0042] In the synchronous mode, the priority rule-based coalescing mechanism 110 may actively send an inquiry, with a coalesced file containing information that has been properly coalesced, to an agent to collect information. To respond, the agent may simply append the requested information at the end of the coalesced file. When the appended coalesced file is returned to the priority rule-based coalescing mechanism 110, the appended information may be coalesced against the information that has been properly coalesced to generate an update coalesced file, which may then be passed onto the next agent.

[0043] In the asynchronous mode, the coalesced file may remain in the priority rule-based coalescing mechanism 110. When information is received by the network communication mechanism 310 from an agent, the coalescing mechanism 330 coalesces the newly received information against the information in the coalesced file to generate an update coalesced file.

[0044] Fig. 5 is an exemplary flowchart for the priority rule-based coalescing mechanism 110, in which information from different agents is coalesced, in a synchronous mode, based on priority rules according to an embodiment of the present invention. The priority rule set-up mechanism 360 first initializes or updates, at act 510, the priority rules and stores them in the priority rule database 365. To perform synchronous coalescing, the coalesced file initialization mechanism 350 constructs, at act 520, an initial coalesced file.

[0045] During the coalescing process, the coalesced file is sent, via the network communication mechanism 310, to each agent, connected to the priority rule-based coalescing



mechanism 110, with an inquiry and the requested information is returned as an appended portion in the coalesced file and sent back to the network communication mechanism 310. The returned coalesced file containing the newly appended information is then coalesced to generate an update coalesced file in which all the information has been properly coalesced. The process continues until all the agents have been enumerated. When this happens, determined at act 530, the coalesced file is stored, at act 535, in the coalesced file database 335.

[0046] During the enumeration, when there is a next agent, determined at act 530, the network communication mechanism 310 sends, at act 540, an inquiry together with the current coalesced file to the next agent. When the network communication mechanism 310 receives, at act 550, returned information (formatted according to, for example, XML), the priority rule access mechanism 370 retrieves, at act 560, the priority rules associated with the next agent. The retrieved priority rules are used to transform, by the transformation mechanism 320 at act 570, the appended information to generate modified information with the pre-defined priorities incorporated. The modified information and the retrieved priority rules are used to coalesce, at act 580, the coalesced file to generate, at act 590, an updated coalesced file. The process then returns to act 530 to determine whether the enumeration continues.

[0047] Fig. 6 is an exemplary flowchart of a process, in which an agent supplies information to the priority rule-based coalescing mechanism 110 in a synchronous mode, according to an embodiment of the present invention. In Fig. 6, an agent receives, at act 610, an inquiry for information from the priority rule-based coalescing mechanism 110. The inquiry may be received with a coalesced file containing, for example, information that has

been properly coalesced. The inquiry may request the agent to collect some information related to certain device such as the timestamp of a file under a file server.

**[0048]** Upon receiving the inquiry, the agent connects, at act 620, to the device and collects, at act 630, the requested information. Before sending the information back to the priority rule-based coalescing mechanism 110, the agent constructs, at act 640, formatted information according to a pre-defined schema or structure and appends, at act 650, the formatted information to the coalesced file. Then, the agent sends, at act 660, the formatted information with the appended coalesced file to the priority rule-based coalescing mechanism 110.

**[0049]** Fig. 7 is an exemplary flowchart for the priority rule-based coalescing mechanism 110, in which information from different sources is coalesced, in an asynchronous mode, based on priority rules according to a different embodiment of the present invention. The priority rule set-up mechanism 360 first initializes or updates, at act 710, the priority rules and stores them in the priority rule database 365. The coalesced file initialization mechanism 350 initializes, at act 720, an initial coalesced file.

**[0050]** In an asynchronous mode, the priority rule-based coalescing mechanism 110 is activated to perform coalescing whenever information is received from an agent. In Fig. 7, when information is received from an agent, determined at act 730, the priority rule-based coalescing mechanism 110 retrieves, at act 750, relevant priority rules from the priority rule database 365. The received information is then, transformed, at act 760, using the retrieved priority rules to generate modified information (that incorporates the priority rules). The modified information is coalesced, at act 770, in relation to the content that is already in the

coalesced file based on the priority rules. An updated coalesced file is generated, at act 780, based on the coalesced result.

[0051] Fig. 8 is an exemplary flowchart of a process, in which an agent supplies information to the priority rule-based coalescing mechanism 110 in an asynchronous mode, according to a different embodiment of the present invention. In an asynchronous mode, a device may actively sends information to its agent. For example, a device may advertise its availability by sending information to its associated agent. An agent receives, at act 810, information from a device in its associated device group. Priori to forwarding the information to the priority rule-based coalescing mechanism 110, the agent constructs, at act 820, formatted information according to some pre-defined syntax or schema. The properly formatted information is then sent, at act 830, to the priority rule-based coalescing mechanism 110.

[0052] While the invention has been described with reference to the certain illustrated embodiments, the words that have been used herein are words of description, rather than words of limitation. Changes may be made, within the purview of the appended claims, without departing from the scope and spirit of the invention in its aspects. Although the invention has been described herein with reference to particular structures, acts, and materials, the invention is not to be limited to the particulars disclosed, but rather extends to all equivalent structures, acts, and, materials, such as are within the scope of the appended claims.